nature structural & molecular biology

Extensive functional overlap between σ factors in *Escherichia coli*

Joseph T Wade[1], Daniel Castro Roa[2], David C Grainger[3], Douglas Hurd[4], Stephen J W Busby[3], Kevin Struhl[1] & Evgeny Nudler[2]

Bacterial core RNA polymerase (RNAP) must associate with a σ factor to recognize promoter sequences. *Escherichia coli* encodes seven σ factors, each believed to be specific for a largely distinct subset of promoters. Using microarrays representing the entire *E. coli* genome, we identify 87 *in vivo* targets of σ^{32}, the heat-shock σ factor, and estimate that there are 120–150 σ^{32} promoters in total. Unexpectedly, 25% of these σ^{32} targets are located within coding regions, suggesting novel regulatory roles for σ^{32}. The majority of σ^{32} promoter targets overlap with those of σ^{70}, the housekeeping σ factor. Furthermore, their DNA sequence motifs are often interdigitated, with RNAPσ^{70} and RNAPσ^{32} initiating transcription *in vitro* with similar efficiency and from identical positions. σ^{E}-regulated promoters also overlap extensively with those for σ^{70}. These results suggest that extensive functional overlap between σ factors is an important phenomenon.

E. coli RNAP consists of five subunits, $\alpha_2\beta\beta'\omega$, that make up the core enzyme. Core RNAP can synthesize RNA, but it must associate with an accessory σ subunit to form RNAP holoenzyme in order to associate with specific DNA sequences located at promoters[1]. RNAP holoenzyme containing σ^{70} ($E\sigma^{70}$), the predominant σ factor in *E. coli*, binds the promoters of housekeeping genes. *E. coli* also encodes six 'alternative' σ factors that allow RNAP holoenzyme to associate with smaller subsets of promoters[2]. Each alternative σ factor is required for expression of certain genes in response to a specific environmental stimulus. Most bacterial genomes encode multiple σ factors that are required for complex cellular processes such as stress response, morphogenesis and virulence[2–5]. Alternative σ factors generally recognize different promoter sequences from the housekeeping σ factor and from one another.

The vast majority of genes are predicted to be transcribed by a single RNAP holoenzyme. In *E. coli*, among genes whose transcription is known to be driven by an alternative σ factor other than σ^{38}, only ~10% are known to be transcribed by σ^{70} holoenzyme[6]. σ^{38} represents an unusual case because the DNA sequence specificity of σ^{38} is very similar to that of σ^{70} (refs. 7,8). Most instances of multiple holoenzymes transcribing a single gene involve binding of different holoenzymes to separate promoter sequences. Hence, the different holoenzymes transcribe messenger RNAs with 5′ UTRs of different lengths. There are very few described examples of different holoenzymes binding overlapping promoter sites in any bacterial species[8–14].

The *E. coli* alternative σ factor σ^{32}, and its homologs from different bacteria, are master regulators of the heat-shock response[3,15]. The general model of the heat-shock response postulates that the increased level of σ^{32} upon heat shock leads to σ switching—that is, substitution of the housekeeping σ^{70} subunit of RNAP with heat shock–specific σ^{32}. σ^{32} holoenzyme ($E\sigma^{32}$) recognizes distinct –10 (CTTGAAA) and –35 (CCCCATNT) promoter elements, thus directing RNAP to heat-shock genes[16]. Traditionally, members of the σ^{32} regulon have been identified using biochemical techniques comparing the abundance of individual transcripts or proteins before and after heat shock[17]. Recently, σ^{32}-regulated genes were identified on a genome-wide scale by virtue of transcriptional changes in response to artificial overproduction of σ^{32} (ref. 18). This first systematic study of the heat-shock regulon revealed at least 26 new heat-shock gene candidates, suggesting that the heat-shock response is far more complex than previously thought.

Here we use chromatin immunoprecipitation coupled with microarrays (ChIP-chip) to identify 87 σ^{32} promoters, many more promoters than have previously been identified. Notably, 25% of the σ^{32} promoters we identify are located within the coding sequences of known genes, suggesting a previously uncharacterized regulatory function for σ^{32}. We demonstrate that two of these promoters drive transcription of a novel class of RNA and that a third drives transcription of an antisense RNA. We also compare the locations of σ^{32} promoters identified in our ChIP-chip study to the locations of σ^{70} promoters identified in a separate study (J.T.W. and K.S., unpublished data). We show that the majority of σ^{32} promoters can also be transcribed by $E\sigma^{70}$ both *in vivo* and *in vitro*, and that $E\sigma^{32}$ and $E\sigma^{70}$ initiate transcription *in vitro* from the same start site at five

[1]Department of Biological Chemistry and Molecular Pharmacology, Harvard University, Boston, Massachusetts 02115, USA. [2]Department of Biochemistry, New York University School of Medicine, New York, New York 10016, USA. [3]School of Biosciences, The University of Birmingham, Edgbaston, Birmingham B15 2TT, UK. [4]Oxford Gene Technology, Begbroke Science Park, Sandy Lane, Yarnton, Oxford OX5 1PF, UK. Correspondence should be addressed to E.N. (evgeny.nudler@med.nyu.edu) or K.S. (kevin@hms.harvard.edu).

Received 16 April; accepted 13 July; published online 6 August 2006; doi:10.1038/nsmb1130

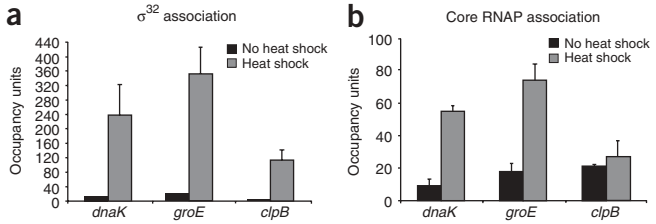

Figure 1 *In vivo* binding of σ^{32} and RNAP to previously identified σ^{32} target promoters. (**a**,**b**) Association of σ^{32} (**a**) and the β subunit of RNAP (**b**) with known σ^{32} targets, before and after heat shock. Occupancy was measured as a background-subtracted ratio of binding of σ^{32} or β to the tested region over binding to a control region located within the coding sequence of *sgrR*. Error bars represent one s.d. from the mean.

of five promoters tested. Furthermore, we show that ~40% of σ^{E} (σ^{24}) promoters are also transcribed by Eσ^{70} *in vivo*, and we demonstrate transcription of two σ^{E} promoters by σ^{70} *in vitro*. Thus, we show that the promoters of two alternative σ factors with apparently distinct core promoter regions overlap extensively with the promoters of σ^{70}. We suggest that extensive overlap between alternative and housekeeping σ factors is a significant phenomenon in *E. coli* and perhaps in other prokaryotic organisms.

RESULTS

Identifying σ^{32} promoters

We used ChIP-chip to directly assess σ^{32} binding across the *E. coli* genome *in vivo*[19–22]. Sites of σ^{32} association must represent σ^{32}-dependent promoters, as σ^{32} does not appreciably associate with elongating RNAP[23]. To validate our ChIP assay, we first determined the association of σ^{32} with three known target promoters, *dnaK*, *groE* and *clpB*, before and after heat shock. We also determined the association of core RNAP (β subunit) in the same cells[22,24] (**Fig. 1**). As expected, there is a large increase in both β and σ^{32} association with all three promoters after heat shock. Notably, σ^{32} is detectable at these promoters before heat shock, demonstrating that a low level of Eσ^{32} can transcribe in normally growing cells. Before heat shock, there is an appreciable level of core RNAP present at each promoter. As the ratio of σ^{32} to β association before heat shock is much lower than that after heat shock, it is likely that much of the β association before heat shock is due to transcription initiation by a non-σ^{32}-containing RNAP holoenzyme.

To identify σ^{32} targets on a genome-wide level in an unbiased manner, we performed a ChIP-chip analysis using microarrays with 60-base oligonucleotides that cover the *E. coli* genome with an average spacing of 223 base pairs (bp)[22]. Cells were harvested after heat shock. Using data from two independent experiments, we identified 87 targets for σ^{32}, with an estimated false discovery rate (FDR) of 1% (see Methods). At a slightly less stringent FDR of 2%, we identified 134 σ^{32} targets. For subsequent analysis, we used the set of 87 targets, although the results are similar with the set of 134 targets. The list of 87 σ^{32} targets is summarized in **Tables 1** and **2**. Raw data are available online (http://www.ogt.co.uk/cms-chip-publications.htm).

For 34 of the 38 targets previously identified by transcriptional profiling[18], we observed σ^{32} association at the same gene, the promoter of the gene or the promoter of a gene located a short distance upstream that is likely to be the upstream gene of an operon. We suspect that the three of the four exceptions (*ydaM*, *htrC* and *cpxP*) are artifacts of the transcriptional profiling experiment, because they ranked near the bottom of the list of candidates. The transcriptional profiling analysis did not identify six putative σ^{32} targets listed by EcoCyc[6], and our ChIP-chip analysis identified only one of these six genes (*rpoD*), suggesting that the remaining five (*ppiC*, *mlc*, *pphA*, *rfaD* and *metA*) are not σ^{32} targets, are bound weakly by σ^{32} or are regulated by σ^{32} only under specific environmental conditions. Of the 65 σ^{32} targets located in promoter regions, the corresponding gene is in the top 10% of genes upregulated by heat shock in 42 cases[25] (**Table 1**). Promoters corresponding to heat shock–induced genes typically show higher levels of σ^{32} association than do other σ^{32} targets.

Many σ^{32} promoters are within known coding sequences

Notably, 22 σ^{32} targets (25%) identified by the ChIP-chip analysis were not located in promoter regions; 20 were assigned to ORFs and 2 to intergenic regions between convergently transcribed genes (**Table 2**). To validate these sites, we used ChIP and quantitative PCR to measure the association of σ^{32} with five targets within ORFs (*sbcD*, *cycA*, *macB*, *dhaM* and *ydeP*) and a target in the intergenic region between the convergently transcribed *tdk* and *ychG* genes (**Fig. 2a**). After heat shock, σ^{32} associated substantially with each of these regions, whereas little or no σ^{32} association was detected before heat shock (**Fig. 2a**), indicating that each of these regions is a genuine σ^{32} target. As the ChIP-chip score (a measure of σ^{32} association derived from the microarray) for the *ydeP* ORF is one of the lowest for any σ^{32} target identified, the number of false positives from our ChIP-chip analysis must be very low, in accord with our FDR estimate of ~1% (see Methods).

Several lines of evidence suggest that many of the ORF-located σ^{32} targets represent transcriptionally active and heat shock–regulated promoters. Two ORF-internal σ^{32} targets (*macB* and *cycA*) are within genes upregulated by overproduction of σ^{32} (ref. 18), and six ORF-internal σ^{32} targets are within the top 10% of genes upregulated after heat shock, which is highly significant (Fisher exact test $P < 0.01$)[25]. It is important to note that these microarray expression studies would be unable to distinguish between full-length mRNAs and shorter RNAs that represent only a fraction of the known gene.

To further investigate the function of Eσ^{32} bound within ORFs, we selected three genes that contained σ^{32} targets, *cycA*, *macB* and *sbcD*. By sequence analysis, we were able to predict the likely binding site for σ^{32} within both *cycA* and *macB*. In each case, the binding site for σ^{32} was oriented in the same direction as the associated ORF. We determined the association of σ^{32} and core RNAP (β subunit) before and after heat shock at three locations within *cycA*, *macB* and *sbcD*: (i) the 5′ end, (ii) close to the predicted σ^{32} binding site (for *sbcD*, the location of the predicted binding site was estimated using the ChIP-chip data) and (iii) downstream of the predicted σ^{32} binding site (**Fig. 2b**). At all three genes, σ^{32} association increased greatly after heat shock, and the association pattern is consistent with binding of σ^{32} to the predicted site (**Fig. 2c**). Notably, very little σ^{32} was associated with the 5′ end of each gene. At *cycA* and *macB*, RNAP association also increased greatly after heat shock, both at the site of σ^{32} association and downstream of this site (**Fig. 2d**). Very little RNAP was associated with the 5′ end of either gene. This strongly suggests that Eσ^{32} is required for heat shock–dependent transcription of RNAs that correspond to a downstream portion of known coding transcripts. It is likely that these shorter transcripts are noncoding, as they lack a suitable Shine-Dalgarno sequence and initiation codon, suggesting a more complex regulatory role. At *sbcD*, RNAP association also increased greatly after heat shock, both at the site of σ^{32} association and upstream of this site (**Fig. 2d**). Very little RNAP was associated with the 3′ end of the *sbcD* gene. This suggests that Eσ^{32} is required

Table 1 σ^{32} targets assigned to promoters

Peak position[a]	ChIP-chip score[b]	Assigned promoter[c]	Heat shock mRNA[d]
12137	111.8	*dnaK*	**58.5**
1338236	72.1	*yciS*	**5.6**
517489	59.9	*ybbN*	**9.9**
1382108	58.0	*ycjX*	**39.3**
3325836	51.0	*ftsJ*	**9.1**
692705	49.9	*ybeZ*	**9.6**
1910676	49.0	*htpX*	**36.1**
4366611	46.8	*fxsA*	**50.7**
4120324	43.4	*hslV*	**16.2**
455885	41.2	*clpP*	**3.3**
661847	37.1	*ybeD*	ND
1860543	34.7	*gapA*	**–3.1**
2732219	31.1	*clpB*	**36.5**
458043	30.3	*lon*	**20.3**
494360	29.1	*htpG*	**33.8**
3643277	28.3	*prlC*	**16.7**
4368697	27.4	*groE*	**77.5**
4435901	27.2	*ytfI*	–0.2
4397378	23.5	*miaA*	**11.9**
3472832	22.4	*rpsL*	ND
2879078	22.0	*ygcI*	**4.8**
1329066	21.0	*topA*	**5.9**
1120249	20.9	*yceP*	**25.5**
3865622	20.4	*ibpA*	**297.4**
3527116	19.7	*yrfG*	**12.1**
3117399	18.3	*yghJ*	ND
2748770	17.8	*grpE*	**24.1**
4570170	16.3	*yjiT*	2.2
4524200	15.9	*yjhI*	**5.5**
1027984	15.7	*yccV*	**34.3**
1894792	15.3	*sdaA*	**23.6**
3210711	14.6	*rpoD*	**7.7**
1744318	13.6	*ydhQ*	**3.7**
3437568	13.2	*yhdN*	**9.5**
2533473	12.6	*crr*	–2.5
2166182	11.7	*b2084*	**4.6**
2925989	10.1	*sdaC*	2.2
1441543	9.0	*ldhA*	**25.9**
3543459	8.7	*yhgH*	1.9
2735126	7.9	*yfiA*	2.3
3766465	7.8	*yibG*	ND
231049	7.6	*yafD*	**6.7**
918375	7.2	*ybjX*	–2.1
3764283	6.9	*yibA*	ND
63622	6.8	*hepA*	**6.8**
239047	6.1	*yafU*	**3.8**
2209265	5.4	*yehR*	3.0
1581861	5.4	*ydeO*	**3.2**
1710454	5.3	*ydgR*	**–6.8**
2217821	4.7	*yehZ*	0.1
4124864	4.7	*rpmE*	**5.7**
3725552	4.7	*yiaA*	**3.7**
1063460	4.7	*yccE*	**5.7**
3427148	4.5	*yrdA*	2.0
2288414	4.4	*narP*	**4.2**
2798757	4.3	*nrdH*	0.8
4482322	4.3	*holC*	**3.8**
1173268	4.2	*mfd*	1.6
1189625	3.9	*phoP*	0.7
4465355	3.9	*treR*	1.5
705196	3.8	*glnS*	–0.7
22199	3.7	*ileS*	–1.4
2520600	3.7	*xapR*	3.0
2771222	3.4	*b2641*	2.6
4538807	3.2	*fimB*	–1.0

[a]Genome coordinate corresponding to the center of the peak microarray probe. [b]Normalized ratio of σ^{32} binding relative to the genomic DNA control; derived from the ChIP-chip analysis. [c]For probes found within nondivergent promoter regions, the gene with the nearest 5′ end to the peak position was chosen. For divergent promoter regions, the gene with the greatest transcriptional induction after heat shock[25] was chosen. Promoters of previously known σ^{32} target genes[18] are underlined. [d]The level of induction ($\log_2$) of mRNA abundance after heat shock for the gene corresponding to the assigned promoter[25]. Values in the top or bottom 10% of this analysis are in bold. ND, not determined.

for heat shock–dependent transcription of an antisense transcript that covers some of the *sbcD* gene.

Determining the Eσ^{32} consensus binding site

Using BioProspector[26], a we defined a two-part consensus motif for σ^{32} (**Fig. 3a**) that is a good match to the previously defined consensus[16,18]. BioProspector identifies conserved sequence motifs, including bipartite motifs, such as the consensus binding site for σ^{32}. The motif we identified is more degenerate than previously identified motifs, perhaps owing to our ability to identify activator-dependent σ^{32} binding sites. At these sites, the affinity of σ^{32} for the promoter in the absence of a transcription activator is likely to be low. The two motif halves were separated by 13, 14 or 15 bp in most cases. This is consistent with a previous study that showed optimum spacing of 15 bp at the *groE* promoter, with 14 bp also being suitable for transcription and 16 bp being incompatible[16].

Determining the overlap of σ^{32} and σ^{70} promoters

In a separate study, we identified σ^{70} targets comprehensively across the *E. coli* genome (J.T.W. and K.S., unpublished data; see **Supplementary Data** online). To compare the positions of σ^{32} target sites with those identified for σ^{70}, we first selected only those σ^{32} targets for which the peak microarray probe corresponded to a position within a promoter that is not adjacent to a divergently transcribed gene (divergently transcribed genes being those whose 5′ ends are adjacent but are transcribed on opposite strands). For each of these 36 σ^{32} targets, we calculated the minimum distance from a σ^{70} target. We also calculated the minimum distance from each of 1,000 randomly selected probe coordinates to a σ^{70} target. The distributions of these distances are shown in **Figure 3b**. There is a highly significant (Fisher exact test $P < 3 \times 10^{-10}$) overlap between the positions of σ^{32} targets and σ^{70} targets. Notably, 56% (20 of 36) of σ^{32} targets analyzed are located within 200 bp of a σ^{70} target, demonstrating that the majority of σ^{32}-regulated promoters are also bound by Eσ^{70}. 56% should be regarded as a lower bound, as σ^{70} targets were only identified under a single growth condition (J.T.W. and K.S., unpublished data). This overlap is particularly dramatic when compared to a model where σ factors bind mutually exclusive sets of promoters. In that case, there would be a substantially lower overlap than that seen with random genome positions. To model this situation, we took advantage of the unbiased transcript map identified in our separate study (J.T.W. and K.S., unpublished data; see **Supplementary Data**). We determined, for 254 transcript 5′ ends (not including transcripts of divergently transcribed genes; transcripts were mapped from cells grown to

Table 2 σ^{32} targets assigned to nonpromoter regions

Peak position[a]	ChIP-chip score[b]	Assigned gene[c]	Heat shock mRNA[d]	Distance from nearest 5′ end (bp)[e]
2780215	17.8	*ypjA*	2.9	662
2385670	15.5	*yfbM/yfbN**	2.5/2.3	714 (*yfbM*)
4429093	10.7	*cycA*	1	913 (*ytfE*)
415417	10.4	*sbcD*	**18.9**	440 (*sbcC*)
921091	10.2	*macB*	**5.2**	722 (*cspD*)
1293531	6.8	*tdk/ychG**	0.5/2.9	708 (*ychG*)
1579583	6.7	*ydeN*	**3.7**	754 (*ydeM*)
2236301	6.5	*mglA*	**6.7**	526 (*mglC*)
2319337	5.7	*atoS*	3.1	551 (*atoC*)
2923838	5.5	*yqcD*	0.8	468
1789863	5.4	*ydiV*	ND	181
1247539	5.1	*dhaM*	**–5.8**	801
2769946	5.0	*yfjU*	1.4	230
4359431	4.7	*cadC*	0.8	526
2764390	4.5	*yfjN*	ND	450
1624219	4.4	*dcp*	–1.4	1,185
516521	3.8	*ybbM*	2	741
1213931	3.8	*ycgF*	**3.4**	649 (*ycgE*)
1609353	3.8	*yneF*	0.4	525
1584068	3.5	*ydeP*	ND	442
2762841	3.4	*yfjL*	1.2	334
3878830	3.3	*recF*	**5.8**	414

[a]Genome coordinate corresponding to the center of the peak microarray probe. [b]Normalized ratio of σ^{32} binding relative to the genomic DNA control; derived from the ChIP-chip analysis. [c]Gene within which the peak probe position was located. For intergenic regions of convergently transcribed genes, both genes are listed (marked with asterisk). Previously identified σ^{32} target genes[18] are underlined. [d]The level of induction ($\log_2$) of mRNA abundance after heat shock for the gene corresponding to the assigned promoter[25]. Values in the top or bottom 10% of this analysis are in bold. ND, not determined. [e]Distance from the peak position to the nearest 5′ end of a gene. This is not always the assigned gene; when it is not, the gene with the nearest 5′ end is listed in parentheses.

exponential phase at 30 °C), the minimum distance to another transcript 5′ end. This simulates the situation for σ factors that bind mutually exclusive sets of promoters, as all transcript 5′ ends are distinct. As a control, we determined, for 254 random genomic coordinates, the minimum distance to a transcript 5′ end. The distributions of these distances are plotted in **Figure 3c**. This simulation demonstrates that any σ^{32} targets located within 300 bp of a σ^{70} target are highly likely (>99%) to represent σ^{32} promoters that can also be transcribed by σ^{70}. (Note that the resolution of this analysis is not sufficient to determine the precise positions of the promoter sequences for each σ factor.)

Close examination of heat-shock promoters reveals many good matches to the –35 and –10 hexamers required for $E\sigma^{70}$ association. Furthermore, the upstream half of the σ^{32} consensus sequence (CTTGAA) is notably similar to the consensus sequence of the σ^{70} –35 hexamer (TTGACA). We aligned the consensus sequence of a σ^{32}-dependent promoter with that of a σ^{70}-dependent promoter so that the overlap of the consensus sequences for the two σ factors was maximized. As the two binding elements for each σ factor can be spaced in three different ways, there are nine potential ways these promoter elements can overlap (**Fig. 3d**). In each case, there is a mismatch of only 1 bp in the upstream promoter elements for σ^{32} and σ^{70}. In five of the nine cases, there is 0 or 1 mismatched bp in the downstream promoter elements for σ^{32} and σ^{70}. This suggests that σ^{32} and σ^{70} can regulate the same promoters by binding overlapping DNA sequences.

The sequences of many σ^{32}-bound promoters contain excellent matches to both the –10 and –35 hexamer consensus sequences for σ^{70}. Quantitative analysis of σ^{32} and σ^{70} association with each of three such promoters, *yciS*, *rpsL* and *hepA*, revealed a large increase in the binding of σ^{32} after heat shock (**Fig. 3e**). σ^{70} associated with each promoter before and after heat shock, although binding decreased about two-fold after heat shock (**Fig. 3f**). Thus, both σ factors are able to bind these promoters, sometimes under the same growth conditions.

Transcription by $E\sigma^{32}$ and $E\sigma^{70}$ *in vitro*

To examine whether $E\sigma^{70}$ can indeed recognize heat-shock promoters, we reconstituted *in vitro* transcription using highly pure *E. coli* RNAP core enzyme and individually purified σ^{70} and σ^{32}. We compared the efficiency of initiation by $E\sigma^{32}$ and $E\sigma^{70}$ at five heat-shock promoters, *groE*, *rpsL*, *clpB*, *htpG* and *lon*, that are known to be regulated by $E\sigma^{32}$. The σ^{70}-specific *T7A1* promoter template was used as a negative control for $E\sigma^{32}$ initiation. Notably, $E\sigma^{70}$ was at least as efficient at each of the heat-shock templates as $E\sigma^{32}$ at either 37 °C or 45 °C (**Fig. 4a**). In contrast, only $E\sigma^{70}$ was able to initiate at the *T7A1* promoter, showing the strict specificity of σ^{32}. This control result also demonstrates the high purity of σ^{32} and σ^{70} (**Supplementary Fig. 1** online) and excludes any potential cross-contamination.

To examine whether the two holoenzymes initiate from exactly the same or different start sites, we modified the transcription reaction to

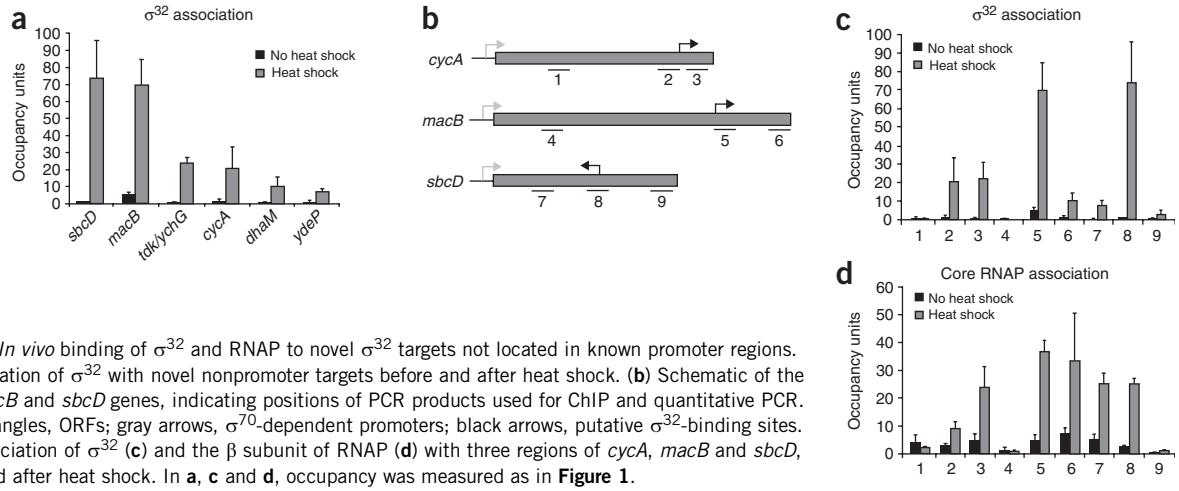

Figure 2 *In vivo* binding of σ^{32} and RNAP to novel σ^{32} targets not located in known promoter regions. (**a**) Association of σ^{32} with novel nonpromoter targets before and after heat shock. (**b**) Schematic of the *cycA*, *macB* and *sbcD* genes, indicating positions of PCR products used for ChIP and quantitative PCR. Gray rectangles, ORFs; gray arrows, σ^{70}-dependent promoters; black arrows, putative σ^{32}-binding sites. (**c**,**d**) Association of σ^{32} (**c**) and the β subunit of RNAP (**d**) with three regions of *cycA*, *macB* and *sbcD*, before and after heat shock. In **a**, **c** and **d**, occupancy was measured as in **Figure 1**.

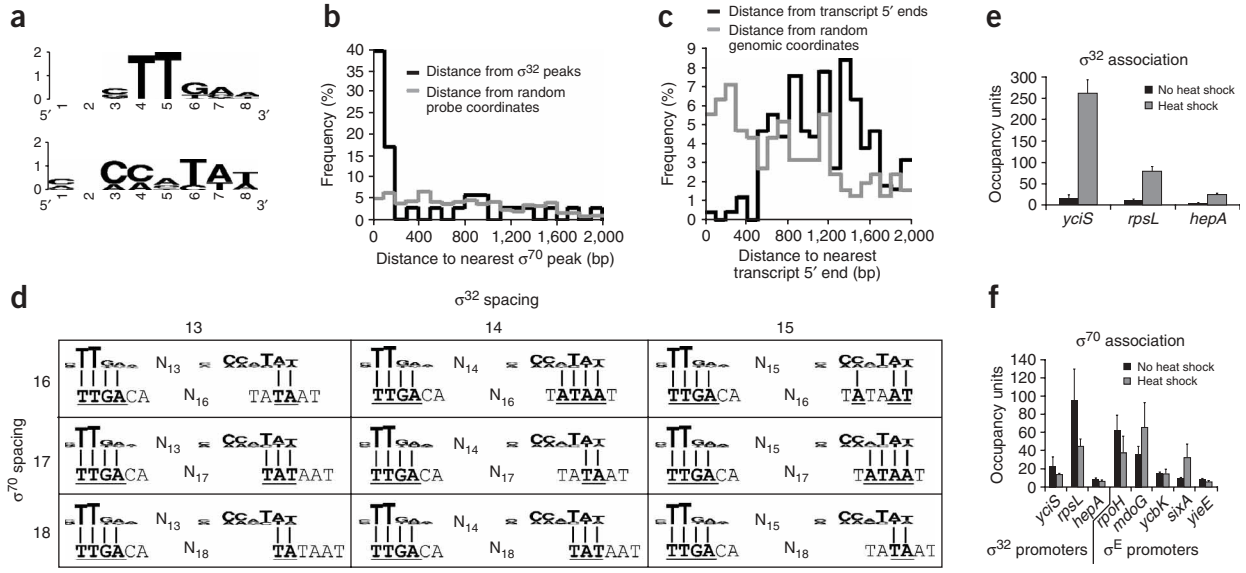

Figure 3 Extensive overlap between σ^{70} targets and targets of σ^{32} and σ^{E}. (**a**) σ^{32} binding motifs identified using BioProspector. (**b**) Distributions of minimum distances to a σ^{70} target from σ^{32} target promoters and from 1,000 random probe positions. (**c**) Distributions of minimum distances between all transcript 5′ ends and transcript 5′ ends of 254 randomly selected, non–divergently transcribed genes, and between all transcript 5′ ends and 254 randomly selected probe positions (see Methods). (**d**) Potential overlap between σ^{32} and σ^{70} promoter elements. Matched bases are indicated by vertical lines and bold, underlined text in the consensus σ^{70} promoter elements. (**e**) Association of σ^{32} with promoters shared by σ^{32} and σ^{70}, before and after heat shock. (**f**) Association of σ^{70} with σ^{32}- and σ^{E}-dependent promoters, before and after heat shock. Occupancy was measured as in **Figure 1**.

allow the formation of promoter-proximal stalled elongation complexes ('walking' transcription assay; see Methods). Three biotin-labeled heat-shock promoter templates, *groE*, *lon* and *dnaK*, were prepared and immobilized on streptavidin beads. Transcription of these immobilized templates was carried out using either $E\sigma^{32}$ or $E\sigma^{70}$, but only three of the four NTPs, so that transcription stalled at a position close to the promoter. In each case, the same predominant RNA products of predicted length were generated, indicating identical transcription start sites for both holoenzymes (**Fig. 4b**). These RNA products belonged to mature elongation complexes, as they were retained on beads after high-salt washing and could be extended upon addition of the missing nucleotide. We repeated the walking transcription assay with the template for *groE*, using different salt conditions (150 mM potassium glutamate) and not including dinucleotides in the transcription reaction (**Supplementary Fig. 2** online). We detected transcripts of the same size as those detected in the previous walking transcription assay (**Fig. 4b**), confirming that our results are not attributable to low-salt conditions or the presence of initiating dinucleotides.

In contrast to our work, an earlier study[27] did not detect transcription of *dnaK* by $E\sigma^{70}$ *in vitro*. We clearly detect transcription of *dnaK* by $E\sigma^{32}$ and $E\sigma^{70}$ from the same start point (**Fig. 4b**). We suggest that this discrepancy could result from differences in the reaction conditions or the method of RNAP holoenzyme and template preparation.

Determining the overlap of σ^{E} and σ^{70} promoters

A combination of transcriptional profiling and bioinformatics has identified a large number of promoters regulated by σ^{E} (σ^{24}), which is required for the envelope stress response[28]. Using the same approach as described above for σ^{32} (**Fig. 3b**), we compared the locations of such σ^{E}-regulated promoters to the targets of σ^{70}. Just as observed for σ^{32}, there is extensive overlap between the positions of σ^{E} targets and those of σ^{70} targets. Indeed, depending on the promoters analyzed

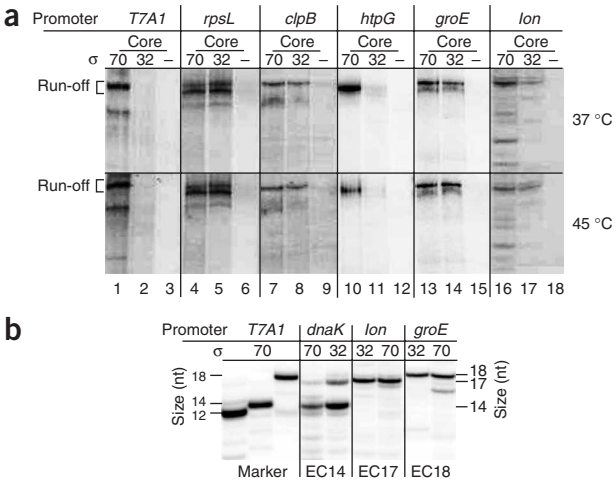

Figure 4 Transcription from heat-shock promoters by $E\sigma^{70}$ and $E\sigma^{32}$ *in vitro*. (**a**) Autoradiogram shows full-size [^{32}P]RNA products (run-off) from reconstituted transcription reactions (see Methods)[40–42,44]. RNAP core was mixed with a three-fold molar excess of σ^{70} or σ^{32} (as indicated) and all four NTP substrates. The reaction was initiated by adding an excess of the DNA template (a PCR fragment with the indicated promoter) and carried out for 5 min at 37 °C or 45 °C. Increasing the concentration of either σ did not result in a greater yield of RNA (data not shown), indicating that both mixtures contained a saturating amount of active σ molecules. The σ^{70}-specific promoter *T7A1* (lanes 1–3) was used as negative control for RNAPσ^{32} and as a reference for RNAPσ^{70} activity. (**b**) Formation of the promoter-proximal elongation complexes (EC) by $E\sigma^{70}$ and $E\sigma^{32}$ on three heat-shock promoter templates (see Methods). Sizes of [^{32}P]RNA products are indicated. Marker was prepared by RNAP walking on *T7A1* template with $E\sigma^{70}$ (ref. 40).

(see Methods), between 31% and 42% of σ^E targets analyzed are located within 300 bp of a σ^{70} target (**Supplementary Fig. 3** online), demonstrating that a substantial proportion of σ^E-regulated promoters can also be bound by $E\sigma^{70}$. This number is a lower bound, for the reasons discussed above with regard to σ^{32}.

We confirmed the binding of σ^{70} to five known σ^E promoters *in vivo* using ChIP (**Fig. 3e**). In each case, we detected σ^{70} at the promoter both before and after heat shock. Notably, σ^{70} binding increased at two promoters, *sixA* and *mdoG*, after heat shock. We also confirmed the association of σ^{70} with two σ^E promoters *in vitro*: $E\sigma^{70}$ transcribed specific products from the *rpoE* and *yfiO* promoters in a run-off transcription assay and from the *rpoE* promoter in a walking transcription assay (**Supplementary Fig. 4** online).

DISCUSSION

ChIP-chip identifies many novel $E\sigma^{32}$ promoters

We have used ChIP-chip to determine the genome-wide binding profile of $E\sigma^{32}$ after heat shock. Our analysis has identified 87 novel targets of $E\sigma^{32}$, including 52 previously undescribed targets. This is far more σ^{32} promoters than expected. We chose a stringent cutoff (estimated FDR of 1%) when analyzing the ChIP-chip data to ensure that almost all called targets are genuine targets of σ^{32}. Our site validation confirms that most, if not all, are genuine. We repeated our analysis using a less stringent cutoff (estimated FDR of 2%). In this case, we identified an additional 47 target regions. Although we have not determined which of these regions are genuinely bound by σ^{32}, we calculated the median distance of each of the 47 targets to the closest gene 5′ end: 152 bp. We performed the same analysis for the 87 high-stringency targets and for 1,000 randomly selected microarray probe positions, yielding average distances of 155 bp and 321 bp, respectively. Thus, the 47 additional σ^{32} targets are significantly more likely to be in promoter regions than are a random set of probe positions (Mann-Whitney $P = 0.005$), and they are positioned similarly to the 87 high-stringency targets, relative to gene 5′ ends. This strongly suggests that most of the 47 low-stringency targets are genuinely bound by σ^{32}. On the basis of this analysis, we estimate that there are 120–150 σ^{32} promoters in *E. coli*, over five times the number of previously identified promoters.

Two studies related to our own[18,29] have used a transcriptomic approach to identify σ^{32}-regulated genes. Their work identified 32 and 49 σ^{32}-dependent promoters, respectively (overlap of 28). One advantage of these studies over our ChIP-chip approach is that they identify not only σ^{32} promoters but also downstream genes within operons that are transcribed by $E\sigma^{32}$. Nevertheless, we identified 87 σ^{32} promoters, many more than both studies combined (overlap of 29 and 37 with the lists of refs. 18 and 29, respectively). ChIP-chip has a number of advantages over transcriptomic approaches for determining the targets of a sequence-specific transcription factor. First, ChIP-chip does not identify genes that are indirectly regulated. Second, ChIP-chip identifies genes that are bound by the protein of interest regardless of whether the level of transcription changes, whereas a transcriptomic analysis requires a significant change in the transcription level. Transcription levels might not change substantially owing to a combination of indirect and direct effects on transcription or because levels were so high to begin with that any changes are undetectable. Third, ChIP-chip does not require overexpression or deletion of the protein of interest, and it can be used to detect binding that is dependent upon specific environmental conditions—for example, conditions required for the presence of other transcription factors.

The σ^{32}-regulated genes identified in ref. 29 are significantly enriched for genes involved in transcription (11 out of 89). We also identified 11 σ^{32}-regulated genes involved in transcription (Fisher's exact $P = 0.03$), and our study and ref. 29 together identify 18. Therefore, σ^{32} regulates at least 18 genes involved in transcription, thus controlling the transcription of many genes indirectly. Notably, we also found that two ribosomal protein genes, *rpsL* and *rpmE*, have σ^{32} promoters. Our analysis of the microarray data of refs. 29 and 25 shows that transcription of ribosomal protein genes as a whole is significantly downregulated by overexpression of σ^{32} (*t*-test $P < 2 \times 10^{-10}$), and by heat shock (*t*-test $P < 3 \times 10^{-15}$; perhaps owing to the effects of σ^{32}). However, transcription of *rpsL* and *rpmE* is upregulated compared to ribosomal protein genes as a whole. The global downregulation of ribosomal protein gene transcription by σ^{32}, coupled with the fact that ribosomal protein genes are highly transcribed in the absence of σ^{32}, explains why the microarray studies[29,25] did not identify these promoters as σ^{32} targets. Our data suggest that σ^{32} is responsible for maintaining the levels of two specific ribosomal protein genes after heat shock, while downregulating the majority.

For many of the σ^{32} promoters identified in this study, no promoter sequence elements were found by the sequence-alignment program BioProspector[26]. We suggest that some of the novel σ^{32} targets identified by our ChIP-chip analysis represent promoters where $E\sigma^{32}$ is recruited by contacts with transcription activators that are functional only after heat shock. Such promoters would lack high-affinity DNA sites for $E\sigma^{32}$, and therefore these sites might not be identified by sequence alignment. It is likely that we cannot detect all promoters that can be transcribed by $E\sigma^{32}$, as our approach would not identify promoters that are bound by σ^{32} under different growth conditions.

Two novel classes of $E\sigma^{32}$ promoter

To our surprise, 25% of the σ^{32} promoters we identified are located within the coding sequences of annotated genes or (in two cases) in intergenic regions between convergently transcribed genes. We have investigated three of these promoters in more detail. The novel σ^{32} promoter within the *sbcD* gene generates an antisense transcript that covers some of the *sbcD* gene. Very few antisense transcripts have been identified previously in *E. coli*[30]. The novel σ^{32} promoters within the *cycA* and *macB* genes generate transcripts that correspond to the downstream portions of the genes themselves. There are a few characterized examples of promoters located within ORFs that drive transcription of downstream genes[6]. However, the transcripts initiated within *cycA* and *macB* are unlikely to encode proteins because they lack a suitable Shine-Dalgarno sequence and initiation codon between the transcription start site and the transcription termination site. Hence, they represent a previously undescribed class of noncoding transcript. To our knowledge, no equivalent transcripts have been identified in other organisms. Notably, in a separate study, we have identified antisense transcripts, ORF-internal transcripts and transcripts in intergenic regions that are likely to be transcribed by $E\sigma^{70}$ (J.T.W. and K.S., unpublished data). These include the transcript initiated within the *macB* gene, presumably because there is sufficient $E\sigma^{32}$ to transcribe this RNA even in the absence of heat shock (**Fig. 2c**). However, the proportion of $E\sigma^{32}$ promoters located within ORFs and between convergently transcribed genes is much higher than that seen for σ^{70} (J.T.W. and K.S., unpublished data).

Extensive overlap of $E\sigma^{32}$ and $E\sigma^{70}$ promoters

By comparing the distribution of $E\sigma^{32}$ identified in this study and $E\sigma^{70}$ identified in a separate study (J.T.W. and K.S., unpublished data) we have shown that there is extensive overlap between promoters bound by $E\sigma^{32}$ and $E\sigma^{70}$ (**Fig. 3**). Many previous studies have identified genes transcribed by multiple holoenzymes. However, the

vast majority of genes are annotated as being transcribed by a single holoenzyme. Indeed, only ~10% of *E. coli* genes annotated as being transcribed by an alternative σ factor, with the exception of σ^{38}, are also annotated as being transcribed by σ^{70} (ref. 6). Thus, the overlap between σ^{32} and σ^{70} is much more extensive than expected. In fact, we show that the majority of σ^{32} promoters analyzed can also be bound by σ^{70} (**Fig. 3b**). This extensive overlap between σ^{32} and σ^{70} targets explains the ability of *rpoH* (σ^{32})-deficient cells to transcribe heat-shock genes[31].

The *rrnB* P1 and *gapA* promoters are known to be transcribed by both $E\sigma^{32}$ and $E\sigma^{70}$ (refs. 12,32) However, no other previously identified σ^{32} promoters that were confirmed by our ChIP-chip analysis have been shown to be transcribed by $E\sigma^{70}$ (ref. 6). Our ChIP-chip analysis did not identify the *rrn* promoters as targets of $E\sigma^{32}$. This is probably due to the transient binding of $E\sigma^{32}$ to these promoters[23]. Notably, $E\sigma^{32}$ and $E\sigma^{70}$ have been found to initiate transcription from the same nucleotide at the *rrnB* P1 promoter[12]. We have also shown this to be the case at all five promoters tested in this study (**Fig. 4**). Hence, we propose that the majority of promoters transcribed by both $E\sigma^{32}$ and $E\sigma^{70}$ contain overlapping binding sites for the two holoenzymes, such that transcription initiates from the same nucleotide. There are very few examples in all bacterial species of different holoenzymes binding overlapping promoter elements[8–14]. This is an elegant mechanism for incorporating additional levels of transcriptional regulation using existing promoters. It is particularly noteworthy given that contacts made between σ^{32} and the –10 hexamer are very different from those made between σ^{70} and the corresponding –10 hexamer[33].

Functional implications of σ factor overlap

By comparing the distributions of previously identified $E\sigma^{E}$-transcribed promoters[28] and $E\sigma^{70}$-transcribed promoters (J.T.W. and K.S., unpublished data), we have shown that there is also extensive overlap between promoters bound by $E\sigma^{E}$ and $E\sigma^{70}$ (**Fig. 3**). It seems unlikely that this results from overlapping promoter elements, as the consensus DNA sites for these σ factors are very different[28]. Notably, ~40% of σ^{E} promoters analyzed are also bound by σ^{70} (**Supplementary Fig. 3**). We have confirmed this overlap at five promoters *in vivo* (**Fig. 3f**) and two promoters *in vitro* (**Supplementary Fig. 4**). Although there are known examples of genes transcribed by both $E\sigma^{E}$ and $E\sigma^{70}$ (ref. 34), the vast majority of σ^{E} promoters have not previously been shown to be transcribed by $E\sigma^{70}$ (ref. 6). Thus, our results demonstrate that extensive overlap of alternative σ factors with σ^{70} is not limited to σ^{32}. We suggest that this is an important feature of transcriptional logic in bacteria and that many alternative σ factors have evolved primarily to augment transcription from σ^{70}-dependent promoters. This would allow for expression of these genes under multiple growth conditions and hence generate more complex transcriptional regulatory patterns.

Despite the apparent importance of overlap between alternative σ factors and σ^{70} in *E. coli*, it seems likely that not all alternative σ factors in bacteria will overlap with a housekeeping σ factor. Some alternative σ factors have highly specialized roles and regulate their target genes only under very specific circumstances; the developmental σ factors in *Bacillus subtilis* are an example[35,36]. In many cases, the genes regulated by these σ factors are specifically required during cellular differentiation and not during vegetative growth. Hence, these genes are unlikely to be transcribed by the housekeeping σ factor. However, it is impossible to rule out the possibility that a substantial number of the targets of these highly specialized alternative σ factors are also targets of the housekeeping σ factor. Other than a handful of isolated examples, we are not aware of any evidence that genes regulated by developmental σ factors are transcriptionally silent in strains where those σ factors are deleted. Additionally, the majority of genes regulated by developmental σ factors are uncharacterized, giving no clue as to their regulation. Finally, there are at least three examples of promoters in *B. subtilis* that are regulated by both the housekeeping σ factor and a developmental σ factor[37–39]. Although the extent of σ sharing in other bacteria remains to be determined, our results indicate that it is important in transcriptional logic.

METHODS

Cell growth. All ChIP experiments used a derivative of *E. coli* strain DY330 that has a TAP tag introduced at the 3′ end of *rpoH*[20]. This strain was cured of λ phage genes by reintroduction of the *bio* locus, and the kan^{R} marker was removed using *flp* recombinase[19]. For heat-shock experiments, cells were grown at 30 °C to an A_{600} of 0.3–0.6. A sample of cells was taken at this point for the 'no heat shock' experiment. The remainder of the cells were then incubated at 43 °C for 5 min in a prewarmed flask. For ChIP-chip experiments, heat shock was performed at 50 °C for 10 min.

***In vitro* transcription.** All templates were obtained by PCR from *E. coli* genomic DNA using Phusion DNA polymerase (Finnzymes) and the following synthetic DNA primers (IDT). *groE*: sense, 5′-ATGTGAGGTGAATCAGGGTTTTCAC-3′; antisense, 5′-TGAGAAAGTCCGTATCTGTTATGGG-3′. *rpsL*: sense, 5′-CAGTGAAGGAGACGAACTGCTAT-3′; antisense, 5′-CGTACGGAATTCACCAGTCAAACCGCGGGCAATA-3′. *lon*: sense, 5′-CATACAATTAGTTAACCAAAA-3′; antisense, 5′-CGTACGGAATTCCAGATGACACGACTGTGCTTC-3′. *htpG*: sense, 5′-CTCCCTTGCCGGGCGTCATAAG-3′; antisense, 5′-CGTACGGAATT CTAGGTCTACCTCAATAATGCCAT-3′. *clpB*: sense, 5′-CCTAAAGAATCAAGACGATCCGG-3′; antisense, 5′-CGTACGGAATTCCAGACGCATAACTCCTCCCATAA-3′. *rpoE*: sense, 5′-GTAGACTTATAATGATAGATAATG-3′; antisense, 5′-CAAATTTCCACGCGCTATCGAAACGCC-3′. *yfiO*: sense, 5′-GCCGTTAACCA GCACTCGCTGGTC-3′; antisense, 5′-GCACAACGAGTACAGCTCACTGC-3′. All templates were purified from low-melting agarose and diluted in TE buffer (10 mM Tris-HCl and 1 mM EDTA) to ~2 pmol μl^{-1}.

E. coli RNAP core enzyme was purified as described[40]. σ^{70} subunit was overexpressed in BL21(DE3) cells (Novagen) carrying pET15*rpoD* plasmid and purified essentially as described[41]. σ^{32} was overexpressed in TOP10 cells (Invitrogen) carrying pET*rpoH* plasmid and purified as described[42].

All transcription reactions in solution (**Fig. 4a**) were carried out in TB50 (50 mM KCl, 10 mM Tris HCl (pH 7.9), 10 mM $MgCl_2$). Core RNAP (0.02 pmol) was mixed with σ^{70} or σ^{32} (0.06 pmol) and NTPs (final concentrations: UTP, ATP and CTP, 250 μM each; GTP, 10 μM) and 2.5 μl of [^{32}P]GTP (3,000 Ci $mmol^{-1}$; Perkin Elmer). The reaction was initiated upon addition of DNA (1 pmol). After 5 min of incubation at 37 °C or 45 °C, the reactions were quenched with 2 volumes of stop solution (20 mM EDTA, 12 M urea and bromophenol blue), then heated at 90 °C for 1 min in a water bath before loading onto 12% PAGE (19:1 acrylamide/bisacrylamide, 7 M urea, 0.5× TBE).

To prepare promoter-proximal stalled elongation complexes (**Fig. 4b**), the transcription reactions were performed in solid phase as follows. Ultra-pure core enzyme (1 pmol) was mixed with σ^{70} or σ^{32} (5 pmol) in the presence of a biotinylated *groE*, *lon* or *dnaK* PCR template (2 pmol) in 20 μl of TB50. After a 5-min incubation at 37 °C, the initiation mixture was added: either 100 μM CpA, 25 μM GTP plus ATP, and 2 μl [^{32}P]CTP (for *groE* to form 18-mer); 100 μM UpApC, 25 μM ATP plus UTP, and 2 μl [^{32}P]GTP (for *lon* to form 17-mer); or 100 μM CpA, 25 μM ATP plus CTP, and 2 μl [^{32}P]GTP (for *dnaK* to form 14-mer). In the *dnaK* template, the point substitution T18C was made to allow one-step formation of a 14-mer complex. All reaction mixtures were incubated for additional 8 min at 37 °C. Afterward, 5 μl of Neutravidin beads (Pierce) were added for 3 min at room temperature, then washed twice with TB700 and three times with TB100. All reaction were quenched with 2 volumes of stop solution and loaded onto 23% PAGE.

Chromatin immunoprecipitation. ChIP was performed as previously described for the β and σ^{70} subunits of RNA polymerase[21]. The protocol was modified slightly for TAP-tagged σ^{32}, which was immunoprecipitated using

IgG-Sepharose (Amersham). IgG binds the protein A moiety of the TAP tag. All experiments were performed in triplicate. ChIP samples were analyzed by quantitative real-time PCR as described[21], using the *sgrR* coding sequence as a normalizing control. Occupancy units represent background-subtracted ratios to the *sgrR* control region. Primer sequences are available on request.

Chromatin immunoprecipitation chip. ChIP-chip was performed in duplicate as described[22] except that total genomic DNA was used as a control. For data collection, we used an Agilent Technologies microarray scanner, and results were extracted using Agilent Technologies image-analysis software with the local background correction option selected. The Cy5:Cy3 intensity ratio was calculated for each spot and plotted against the corresponding position on the *E. coli* MG1655 chromosome. An arbitrary cutoff was chosen (3.0), and all probes that had an intensity ratio greater than this value in both experiments were selected as targets. Adjacent target probes were merged, the target position being defined by the center of the probe with the highest average intensity ratio. To estimate the number of false positives, one data set was randomized with respect to probe position. The number of targets was then recalculated. This was repeated 50 times. On average, 2 targets were found with randomized data sets, compared to 180 (merged to 87) with real data sets. Thus, we estimate the FDR to be ~1% for our high-stringency set of 87 targets. If the estimated FDR was raised to 2%, 47 more target regions were identified.

ChIP-chip targets were defined as promoter or ORF by determining the position of the target probe relative to known genes. If the target probe was located in a promoter, it was defined as a promoter target. For promoters that are divergently transcribed, the target was associated with the gene that shows the highest transcriptional upregulation after heat shock[25]. If the target probe was located in an ORF and there were no probes between the nearest 5′ gene end and the probe, it was defined as a promoter target and was associated with the promoter of that gene. If the target probe was located in an ORF and was flanked on both sides by probes within the same ORF, it was defined as an ORF target. Note that the number of promoter targets is likely to be an overestimate, because some targets in ORFs may be incorrectly defined as promoter targets. Conserved sequence elements were identified using BioProspector[26].

Comparison of binding sites for different σ factors. We selected those σ^{32} ChIP-chip targets that were located within known promoter regions of genes that are not divergently transcribed from an adjacent gene. We then compared the locations of these targets to the locations of σ^{70} targets identified in a separate study (J.T.W. and K.S., unpublished data). For each σ^{32} target, we calculated the distance to the nearest σ^{70} target. We also generated a list of 10,000 random genomic coordinates and compared the location of each random coordinate to the locations of σ^{70} targets. For each random coordinate, we calculated the distance to the nearest σ^{70} target.

To compare σ^{E} and σ^{70} target locations, we selected σ^{E} promoters identified in ref. 28. We selected only promoters of genes that are not divergently transcribed from an adjacent gene. We chose either the genomic coordinate of the center of the σ^{E}-binding site or, when this was unknown, a position 106 bp upstream of the start codon of the corresponding gene. The latter is the average distance between known σ^{E}-binding sites and the ATG of the corresponding gene. We then compared the locations of these σ^{E} binding sites in the same way as for σ^{32} targets. Ref. 28 describes three classes of promoters in the σ^{E} regulon: (i) promoters of genes significantly upregulated by σ^{E} in a microarray analysis, where the 5′ end of a σ^{E}-dependent transcript could be mapped; (ii) promoters of genes significantly upregulated or downregulated by σ^{E} in a microarray analysis, where the 5′ end of a σ^{E}-dependent transcript could not be mapped; and (iii) promoters of genes not significantly upregulated by σ^{E} in a microarray analysis, where the 5′ end of a σ^{E}-dependent transcript could be mapped. Depending on the classes of genes used in the analysis, the percentage of σ^{E} promoters within 300 bp of a σ^{70} promoter varies from 31% to 48%. Genome coordinates used in this analysis are shown in **Supplementary Table 1** online.

Note that the antibody to σ^{70} cross-reacts in western blots with the flagellar σ factor FliA[43]. Hence, it is possible that a small fraction of the targets identified for $E\sigma^{70}$ are in fact targets of holoenzyme containing FliA. However, this is likely to be a very small fraction, as there are believed to be very few promoters bound by FliA (currently only 15 annotated promoters[6]).

Simulating the distributions of promoters for σ factors that bind completely distinct sets of promoters. We determined an unbiased transcript map in a separate study (J.T.W. and K.S., unpublished data; see **Supplementary Data**). To simulate the distances between promoters for two σ factors that bind completely distinct subsets of promoters, we determined, for 254 transcript 5′ ends (not including transcripts of divergently transcribed genes), the minimum distance to another transcript 5′ end (including divergently transcribed genes). As a control we determined, for 254 random genomic coordinates, the minimum distance to a transcript 5′ end. Genome coordinates used in this analysis are shown in **Supplementary Table 1**.

Comparison of σ^{32} target positions with gene starts. We determined the minimum distance between gene starts and each of (i) the coordinates of the 87 high-stringency σ^{32} targets, (ii) the coordinates of the 47 additional σ^{32} targets when the FDR was set to 2%, and (iii) 1,000 randomly selected probe coordinates.

Raw chromatin immunoprecipitation chip data. All raw data files can be downloaded from the following website: http://www.ogt.co.uk/cms-chip-publications.htm. These text files can be viewed with a ChIP-chip browser that can be downloaded from the same website.

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS

We thank G. Butland, J. Greenblatt and A. Emili (University of Toronto) for the TAP-tagged *rpoH* strain, N. Reppas, R.J. Geisberg, Z. Moqtaderi, M. Schwabish, X. Fan and L.T. Westblade for helpful discussions and C. Baisden for expert technical assistance. We also thank C. Gross for communicating results before publication. J.T.W. was supported by a Charles A. King Trust Postdoctoral Fellowship, Bank of America, Co-Trustee. This work was supported by US National Institutes of Health grants GM30186 (K.S.), and GM72814 and GM58750 (E.N.).

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Published online at http://www.nature.com/nsmb/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Gross, C.A., Lonetto, M. & Losick, R. Bacterial sigma factors. in *Transcriptional Regulation* Vol. 1 (eds. Knight, S.L. & Yamamoto, K.R.) 129–176 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA, 1992).
2. Gruber, T. & Gross, C.A. Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu. Rev. Microbiol.* **57**, 441–466 (2003).
3. Segal, G. & Ron, E.Z. Regulation of heat-shock response in bacteria. *Ann. NY Acad. Sci.* **851**, 147–151 (1998).
4. Kazmierczak, M.J., Wiedmann, M. & Boor, K.J. Alternative sigma factors and their roles in bacterial virulence. *Microbiol. Mol. Biol. Rev.* **69**, 527–543 (2005).
5. Losick, R. RNA polymerase sigma factors and cell-specific gene transcription in a simple developing organism. *Harvey Lect.* **90**, 1–17 (1994–1995).
6. Keseler, I.M. *et al.* EcoCyc: a comprehensive database resource for *Escherichia coli*. *Nucleic Acids Res.* **33**, D334–D337 (2005).
7. Gaal, T. *et al.* Promoter recognition and discrimination by ESigmaS RNA polymerase. *Mol. Microbiol.* **42**, 939–954 (2001).
8. Hengge-Aronis, R. Stationary phase gene regulation: what makes an *Escherichia coli* promoter SigmaS-selective? *Curr. Opin. Microbiol.* **5**, 591–595 (2002).
9. Dupuy, B. *et al.* Regulation of toxin and bacteriocin gene expression in *Clostridium* by interchangeable RNA polymerase sigma factors. *Mol. Microbiol.* **60**, 1044–1057 (2006).
10. Dominguez-Cuevas, P., Marin, P., Ramos, J.L. & Marques, S. RNA polymerase holoenzymes can share a single transcription start site for the Pm promoter. Critical nucleotides in the -7 to -18 region are needed to select between RNA polymerase with Sigma38 or Sigma32. *J. Biol. Chem.* **280**, 41315–41323 (2005).
11. Johnson, W.C., Moran, C.P. & Losick, R. Two RNA polymerase sigma factors from *Bacillus subtilis* discriminate between overlapping promoters for a developmentally regulated gene. *Nature* **302**, 800–804 (1983).
12. Newlands, J.T., Gaal, T., Mecsas, J. & Gourse, R.L. Transcription of the *Escherichia coli rrnB* P1 promoter by the heat shock RNA polymerase (E sigma 32) *in vitro*. *J. Bacteriol.* **175**, 661–668 (1993).
13. Wang, P.Z. & Doi, R.H. Overlapping promoters transcribed by *Bacillus subtilis* Sigma 55 and Sigma 37 RNA polymerase holoenzymes during growth and stationary phases. *J. Biol. Chem.* **259**, 8619–8625 (1984).

14. Tatti, K.M. & Moran, C.P. Utilization of one promoter by two forms of RNA polymerase from *Bacillus subtilis*. *Nature* **314**, 190–192 (1985).
15. Grossman, A.D., Straus, D.B., Walter, W.A. & Gross, C.A. Sigma 32 synthesis can regulate the synthesis of heat shock proteins in *Escherichia coli*. *Genes Dev.* **1**, 179–184 (1987).
16. Wang, Y. & deHaseth, P.L. Sigma 32-dependent promoter activity in vivo: sequence determinants of the *groE* promoter. *J. Bacteriol.* **185**, 5800–5806 (2003).
17. Lemaux, P.G., Herendeen, S.L., Bloch, P.L. & Neidhardt, F.C. Transient rates of synthesis of individual polypeptides in *E. coli* following temperature shifts. *Cell* **13**, 427–434 (1978).
18. Zhao, K., Liu, M. & Burgess, R.R. The global transcriptional response of *Escherichia coli* to induced Sigma 32 protein involves Sigma 32 regulon activation followed by inactivation and degradation of Sigma 32 *in vivo*. *J. Biol. Chem.* **280**, 17758–17768 (2005).
19. Datsenko, K.A. & Wanner, B.L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. USA* **97**, 6640–6645 (2000).
20. Butland, G. *et al.* Interaction network containing conserved and essential protein complexes in *Escherichia coli*. *Nature* **433**, 531–537 (2005).
21. Wade, J.T., Reppas, N.B., Church, G.M. & Struhl, K. Genomic analysis of LexA binding reveals the permissive nature of the *Escherichia coli* genome and identifies unconventional target sites. *Genes Dev.* **19**, 2619–2630 (2005).
22. Grainger, D.C., Hurd, D., Harrison, M., Holdstock, J. & Busby, S.J. Studies of the distribution of *Escherichia coli* cAMP-receptor protein and RNA polymerase along the *E. coli* chromosome. *Proc. Natl. Acad. Sci. USA* **102**, 17693–17698 (2005).
23. Raffaelle, M., Kanin, E.I., Vogt, J., Burgess, R.R. & Ansari, A.Z. Holoenzyme switching and stochastic release of Sigma factors from RNA polymerase *in vivo*. *Mol. Cell* **20**, 357–366 (2005).
24. Wade, J.T. & Struhl, K. Association of RNA polymerase with transcribed regions in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **101**, 17777–17782 (2004).
25. Richmond, C.S., Glasner, J.D., Mau, R., Jin, H. & Blattner, F.R. Genome-wide expression profiling in *Escherichia coli* K-12. *Nucleic Acids Res.* **27**, 3821–3835 (1999).
26. Liu, Y. *et al.* A suite of web-based programs to search for transcriptional regulatory motifs. *Nucleic Acids Res.* **32**, W204–W207 (2004).
27. Cowing, D.W. *et al.* Consensus sequence for *Escherichia coli* heat shock gene promoters. *Proc. Natl. Acad. Sci. USA* **82**, 2679–2683 (1985).
28. Rhodius, V.A., Suh, W.C., Nonaka, G., West, J. & Gross, C.A. Conserved and variable functions of the SigmaE stress response in related genomes. *PLoS Biol.* **4**, e2 (2006).
29. Nonaka, G., Blankschien, M., Herman, C., Gross, C.A. & Rhodius, V.A. Regulon and promoter analysis of the *E. coli* heat shock factor, Sigma 32, reveals a multifaceted cellular response to heat stress. *Genes Dev.* **20**, 1776–1789 (2006).
30. Kawano, M., Reynolds, A.A., Miranda-Rios, J. & Storz, G. Detection of 5′- and 3′-UTR-derived small RNAs and *cis*-encoded antisense RNAs in *Escherichia coli*. *Nucleic Acids Res.* **33**, 1040–1050 (2005).
31. Zhou, Y.N., Kusukawa, N., Erickson, J.W., Gross, C.A. & Yura, T. Isolation and characterization of *Escherichia coli* mutants that lack the heat shock sigma factor Sigma 32. *J. Bacteriol.* **170**, 3640–3649 (1988).
32. Charpentier, B. & Branlant, C. The *Escherichia coli gapA* gene is transcribed by the vegetative RNA polymerase holoenzyme E Sigma 70 and by the heat shock RNA polymerase E Sigma 32. *J. Bacteriol.* **176**, 830–839 (1994).
33. Kourennaia, O.V., Tsujikawa, L. & deHaseth, P.L. Mutational analysis of *Escherichia coli* heat shock transcription factor Sigma 32 reveals similarities with Sigma 70 in recognition of the -35 promoter element and differences in promoter DNA melting and -10 recognition. *J. Bacteriol.* **187**, 6762–6769 (2005).
34. Erickson, J.W., Vaughn, V., Walter, W.A., Neidhardt, F.C. & Gross, C.A. Regulation of the promoters and transcripts of *rpoH*, the *Escherichia coli* heat shock regulatory gene. *Genes Dev.* **1**, 419–432 (1987).
35. Eichenberger, P. *et al.* The program of gene transcription for a single differentiating cell type during sporulation in *Bacillus subtilis*. *PLoS Biol.* **2**, e328 (2004).
36. Wang, S.T. *et al.* The forespore line of gene expression in *Bacillus subtilis*. *J. Mol. Biol.* **358**, 16–37 (2006).
37. Popham, D.L. & Setlow, P. Cloning, nucleotide sequence, and regulation of the *Bacillus subtilis pbpF* gene, which codes for a putative class A high-molecular-weight penicillin-binding protein. *J. Bacteriol.* **175**, 4870–4876 (1993).
38. Ramirez, M.I., Castellanos-Juarez, F.X., Yasbin, R.E. & Pedraza-Reyes, M. The *ytkD* (*mutTA*) gene of *Bacillus subtilis* encodes a functional antimutator 8-Oxo-(dGTP/GTP)ase and is under dual control of Sigma A and Sigma F RNA polymerases. *J. Bacteriol.* **186**, 1050–1059 (2004).
39. Paul, S., Zhang, X. & Hulett, F.M. Autoinduction of *Bacillus subtilis phoPR* operon transcription results from enhanced transcription from EsigmaA- and EsigmaE-responsive promoters by phosphorylated PhoP. *J. Bacteriol.* **186**, 4262–4275 (2004).
40. Nudler, E., Gusarov, I. & Bar-Nahum, G. Methods of walking with the RNA polymerase. *Methods Enzymol.* **371**, 160–169 (2003).
41. Borukhov, S. & Goldfarb, A. Recombinant *Escherichia coli* RNA polymerase: purification of individually overexpressed subunits and in vitro assembly. *Protein Expr. Purif.* **4**, 503–511 (1993).
42. Nguyen, L.H., Jensen, D.B. & Burgess, R.R. Overproduction and purification of Sigma 32, the *Escherichia coli* heat shock transcription factor. *Protein Expr. Purif.* **4**, 425–433 (1993).
43. Breyer, M.J., Thompson, N.E. & Burgess, R.R. Identification of the epitope for a highly cross-reactive monoclonal antibody on the major Sigma factor of bacterial RNA polymerase. *J. Bacteriol.* **179**, 1404–1408 (1997).
44. Bar-Nahum, G. & Nudler, E. Isolation and characterization of Sigma70-retaining transcription elongation complexes from *Escherichia coli*. *Cell* **106**, 443–451 (2001).